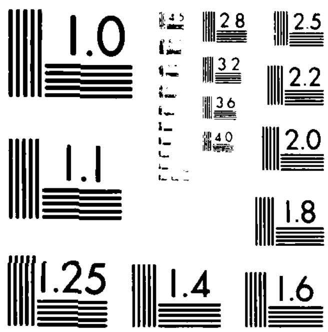
AD-A149 192 RELATIVE-ENTROPY MINIMIZATION WITH UNCERTAIN CONSTRAINTS--THEORY AND APPLICATION TO SPECTRUM ANALYSIS(U) NAVAL RESEARCH LAB WASHINGTON DC 1/1

UNCLASSIFIED R W JOHNSON ET AL. 31 DEC 84 NRL-MR-5478 F/G 12/1 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

(2)

NRL Memorandum Report 5478

Relative-Entropy Minimization with Uncertain Constraints — Theory and Application to Spectrum Analysis

R. W. JOHNSON AND J. E. SHORE

Computer Science and Systems Branch
Information Technology Division

December 31, 1984

NAVAL RESEARCH LABORATORY
Washington, D.C.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NRL Memorandum Report 5478	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Research Laboratory 6b. OFFICE SYMBOL (If applicable) Code 7590	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Washington, DC 20375-5000	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: (See page ii) PROJECT NO: TASK NO.: WORK UNIT ACCESSION NO: DN980-167

11. TITLE (Include Security Classification)

Relative-Entropy Minimization with Uncertain Constraints — Theory and Application to Spectrum Analysis

12. PERSONAL AUTHOR(S)

Johnson, R.W. and Shore, J.E.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM ______ TO ______	1984 December 31	17

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Spectral analysis; Spectrum analysis; Maximum entropy; Cross entropy (Continues)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The relative-entropy principle ("principle of minimum cross entropy") is a provably optimal information-theoretic method for inferring a probability density from an initial ("prior") estimate together with constraint information that confines the density to a specified convex set. Typically the constraint information takes the form of linear equations that specify the expectation values of given functions. This paper discusses the effect of replacing such linear-equality constraints with quadratic constraints that require linear constraints to hold approximately, to within a specified error bound. The results are applied to the derivation of a new multisignal spectrum-analysis method that simultaneously estimates a number of power spectra given: (1) an initial estimate of each; (2) imprecise values of the autocorrelation function of their sum; and (3) estimates of the error in measurement of the autocorrelation values. One application is to separate estimation of the spectra of a signal and independent additive noise, based on imprecise measurements of the autocorrelations of the signal plus noise. The new method is an extension of multisignal relative-entropy spectrum analysis (with exact autocorrelations). The two methods are compared, and connections with previous related work are indicated.

(Continues)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
R. W. Johnson	(202) 767-3490	Code 7590

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.
61153N		RR014-09-41
61153N		RR021-05-42

18. SUBJECT TERMS (Continued)

Relative entropy
Information theory
Prior estimates
Initial estimates

19. ABSTRACT (Continued)

Mathematical properties of the new method are discussed, and an illustrative numerical example is presented.

CONTENTS

RELATIVE-ENTROPY MINIMIZATION WITH UNCERTAIN CONSTRAINTS — THEORY AND APPLICATION TO SPECTRUM ANALYSIS

I. INTRODUCTION

The relative-entropy principle (REP) is a general, information-theoretic method for inference when information about an unknown probability density $q^\dagger$ consists of an *initial* estimate p and additional *constraint* information that restricts $q^\dagger$ to a specified convex set of probability densities. Typically the constraint information consists of linear-equality constraints—expected values

$$\bar{f}_r = \int f_r(x) q^\dagger(x)\, dx \tag{1}$$

for known $f_r(x)$ and $\bar{f}_r$, $r = 0, 1, \ldots, M$. The principle states that one should choose the *final* estimate q that satisfies

$$H(q, p) = \min_{q'} H(q', p),$$

where H is the *relative entropy* (cross entropy, discrimination information, directed divergence, I-divergence, K-L number, etc.),

$$H(q, p) = \int q(x) \log \frac{q(x)}{p(x)}\, dx, \tag{2}$$

and where q' varies over the set of densities that satisfy the constraints. When these are linear-equality constraints (1), the final estimate has the form

$$q(x) = p(x) \exp\left[-\alpha - \sum_r \beta_r f_r(x)\right], \tag{3}$$

where the β_r and α are Lagrangian multipliers determined by (1) (with $q^\dagger$ replaced by q) and by the normalization constraint

$$\int q(x)\, dx = 1. \tag{4}$$

Properties of REP solutions and conditions for their existence are discussed in [1, 2]. Expressed in terms of the expected values and the Lagrangian multipliers, the relative entropy at the minimum is given by

$$H(q, p) = -\alpha - \sum_r \beta_r \bar{f}_r. \tag{5}$$

The normalization multiplier α is given by

$$\alpha = \log \int p(x) \exp\left[-\sum_r \beta_r f_r(x)\right] dx. \tag{6}$$

The quantity $Z = e^\alpha$ is often referred to as the partition function. If the partition function can be evaluated analytically—*i.e.*, if the integral in (6) can be performed—the relations

$$-\frac{\partial \alpha}{\partial \beta_r} = \bar{f}_r \tag{7}$$

Manuscript approved August 30, 1984.

can sometimes be solved to express the β_r as functions of the expected values $\overline{f}_r$. If not, various computational methods can be used to find the values for the α and β_r in (3) that satisfy (1) and (4) [3]. As a general method of statistical inference, the REP was first introduced by Kullback [4], has been advocated in various forms by others [5, 6, 7], and has been applied in a variety of fields (for a list of references, see [3]).

Informally speaking, of the densities that satisfy the constraints, the REP selects the one that is closest to p in the sense measured by relative entropy. In more formal terms, the REP can be justified on the basis of the information-theoretic properties of relative entropy [4], or on the basis of consistency axioms for logical inference [8]. In applications of the REP, the known expected values $\overline{f}_r$ in (1) frequently correspond to physical measurements. Such measurements usually are subject to error so that strict equality in (1) is unrealistic.

In the next section we discuss the REP with "uncertain constraints," a form of the principle appropriate for applications with uncertainty in the expected values. In the third section, relative-entropy minimization with uncertain constraints is applied to spectrum analysis; a relative-entropy spectrum estimate from uncertain autocorrelations is derived. The fourth and fifth sections are devoted to a numerical example and a concluding discussion, respectively.

II. RELATIVE-ENTROPY MINIMIZATION WITH UNCERTAIN CONSTRAINTS

In this Section we extend the results on the REP with linear-equality constraints to incorporate uncertainty about the values of the $\overline{f}_r$ in (1). We define an error vector $\boldsymbol{v}$ with components

$$v_r = \int f_r(x) q^\dagger(x)\, dx - \overline{f}_r . \tag{8}$$

A simple generalization would be to replace the set of constraints (1) with a bound on the magnitude of $\boldsymbol{v}$:

$$\sum_r \left(\int f_r(x) q^\dagger(x)\, dx - \overline{f}_r \right)^2 \le \varepsilon^2 . \tag{9}$$

However, all components v_r may not have equal uncertainty, and different components may be correlated. We therefore replace (9) with the more general constraint

$$\sum_{rs} M_{rs} v_r v_s \le \varepsilon^2 . \tag{10}$$

In matrix notation this is

$$\boldsymbol{v} \cdot \mathbf{M} \boldsymbol{v} \le \varepsilon^2 , \tag{11}$$

where $\mathbf{M}$ is any positive-definite matrix.

We assume that we are given an initial estimate p of $q^\dagger$, measured values $\overline{f}_r$ of the expectations (1) of functions f_r for a finite set of indices r, and an error estimate ε. We will first derive the form of the final estimate q under the assumption that the constraint has the form (9) and that the $\overline{f}_r$ are 0; that is, we assume a constraint

$$\sum_r \left(\int f_r(x) q^\dagger(x)\, dx \right)^2 \le \varepsilon^2 . \tag{12}$$

Next we show how to reduce the more general constraint (10) to this case. We conclude this section with a remark on the relation between the result with $\varepsilon > 0$ and that for "exact constraints" ($\varepsilon = 0$).

Our problem is to minimize the relative entropy $H(q,p)$ subject to the constraint (12) (with q in place of q^*) and the normalization constraint (4). If the initial estimate satisfies the constraint (i.e. (12) holds with p in place of q), then setting $q = p$ gives the minimum. Otherwise equality holds in (12), and the criterion for a minimum is that the variation of

$$\int q(x) \log \frac{q(x)}{p(x)} dx + \lambda \sum_r \left[\int f_r(x) q(x) \, dx \right]^2 + (\alpha - 1) \int q(x) \, dx \qquad (13)$$

with respect to $q(x)$ is zero for some Lagrange multipliers $\lambda > 0$, corresponding to (12), and $\alpha - 1$, corresponding to (4). (We write $\alpha - 1$ instead of α for later convenience.) With $\lambda > 0$, the criterion intuitively implies that a small change δq in q that leaves $\int q(x) \, dx$ fixed and decreases $H(q,p)$ must increase the error term $\sum_r \left[\int f_r(x) q(x) \, dx \right]^2$.

Equating the variation of (13) to zero gives

$$\log \frac{q(x)}{p(x)} + \alpha + \lambda \sum_r 2 f_r(x) \int f_r(x') q(x') \, dx' = 0.$$

Therefore q satisfies

$$q(x) = p(x) \exp \left[-\alpha - \sum_r \beta_r f_r(x) \right] \qquad (14)$$

where

$$\beta_r = 2\lambda \int f_r(x) q(x) \, dx. \qquad (15)$$

Conversely, if q has the form (14), and if α, λ, and the β_r are chosen so that (15), the constraint (12), and the normalization condition (4) hold, then q is a solution to the minimization problem. But if (15) holds, the constraint with equality is equivalent to

$$\sum_r \left(\frac{\beta_r}{2\lambda} \right)^2 = \varepsilon^2,$$

or to

$$\lambda = \frac{1}{2\varepsilon} \|\boldsymbol{\beta}\|,$$

where we have written $\|\boldsymbol{\beta}\|$ for the Euclidean norm $(\sum_r \beta_r^2)^{1/2}$. Thus if we choose α and β_r in (14) so that (4) and

$$\varepsilon \frac{\beta_r}{\|\boldsymbol{\beta}\|} = \int f_r(x) q(x) \, dx \qquad (16)$$

hold, then the constraint (12) will be satisfied, and we can ensure that (15) holds by the choice of λ.

Next assume a constraint of the general form (10), (8), with a symmetric, positive-definite matrix $\mathbf{M}$. Then there is a matrix $\mathbf{A}$, not in general unique, such that $\mathbf{A}^t \mathbf{A} = \mathbf{M}$. Now

$$\boldsymbol{v} \cdot \mathbf{M} \boldsymbol{v} = \boldsymbol{v} \cdot \mathbf{A}^t \mathbf{A} \boldsymbol{v} = (\mathbf{A}\boldsymbol{v}) \cdot (\mathbf{A}\boldsymbol{v}) = \sum_r \left(\sum_s A_{rs} v_s \right)^2,$$

and so the constraint assumes the form

$$\sum_r u_r^2 \le \varepsilon^2, \qquad (17)$$

where

$$u_r = \sum_s A_{rs} v_s$$

In view of (4) we may rewrite (8) as

$$v_r = \int (f_r(x) - \bar{f}_r) q(x)\, dx$$

and obtain

$$u_r = \int \sum_s A_{rs} (f_s(x) - \bar{f}_s) q(x)\, dx$$

Defining

$$g_r(x) = \sum_s A_{rs} (f_s(x) - \bar{f}_s), \tag{18}$$

we obtain

$$\sum_r \left[\int g_r(x) q(x)\, dx \right]^2 \leq \varepsilon^2 \tag{19}$$

from (17). Thus constraints of the general form (10) can be transformed to (19), which is of the same form as (12).

We note that (14) is identical to (3): the functional form of the solution with uncertain constraints is the same as that for exact constraints. The difference is that, for uncertain constraints, the conditions that determine the β_r have the general form (16). These conditions reduce to the exact-constraint case for $\varepsilon = 0$. One way of viewing this identity of form for the solutions of the two problems is to note that every solution q of an uncertain-constraint problem is simultaneously a solution of an exact-constraint problem with the same functions f_k and appropriately modified values for the $\bar{f}_k$.

The relative entropy at the minimum may be computed by substituting (14) into (2), which leads to

$$H(q, p) = -\alpha - \sum_r \beta_r \int f_r q(x)\, dx. \tag{20}$$

In the case of non-zero expected values, $\bar{f}_r \neq 0$, (16) becomes

$$\varepsilon \frac{\beta_r}{\|\boldsymbol{\beta}\|} = \int f_r(x) q(x)\, dx - \bar{f}_r. \tag{21}$$

(For simplicity we take $\mathbf{M}$ to be the identity.) Substituting (21) into (20) yields

$$H(q, p) = -\alpha - \sum_r \beta_r \bar{f}_r - \varepsilon \|\boldsymbol{\beta}\|, \tag{22}$$

which is the generalization of (5) in the case of uncertain constraints. The normalization multiplier α has the same functional form as in the exact-constraint case (6); the generalization of (7) therefore results from differentiating (6), which yields

$$-\frac{\partial \alpha}{\partial \beta_r} = \int f_r(x) q(x)\, dx,$$

and then substituting (21), which yields

$$-\frac{\partial \alpha}{\partial \beta_r} = \bar{f}_r + \varepsilon \frac{\beta_r}{\|\boldsymbol{\beta}\|}. \tag{23}$$

Note that (22) and (23) reduce respectively to (5) and (7) when $\varepsilon = 0$.

III. APPLICATION TO SPECTRUM ANALYSIS

Relative-Entropy Spectrum Analysis (RESA) is an extension of Burg's Maximum-Entropy Spectral Analysis (MESA) [9, 10] that was introduced by Shore [11]. Like MESA, it estimates a spectrum from values of the autocorrelation function. RESA, however, also takes into account prior information in the form of an initial estimate of the spectrum. Multisignal RESA (MRESA), introduced by Shore and Johnson [12], simultaneously estimates the power spectra of several signals when an initial estimate for each spectrum is available and new information is obtained in the form of values of the autocorrelation function of the sum. The resulting *final* estimates are the solution of a constrained minimization problem: they are consistent with the autocorrelation information and otherwise as similar as possible to the respective initial estimates in a precisely defined information-theoretic sense. MRESA has recently been extended by Johnson, Shore, and Burg to incorporate weighting factors associated with each initial spectrum estimate to allow for the fact that initial estimates may not be uniformly reliable [13].

The autocorrelation values were treated in [11, 12, 13] as exactly given. Usually, however, these are estimated or measured values subject to error. By basing a derivation on the REP with uncertain constraints, we will show how to incorporate an error bound to allow for uncertainty in autocorrelation values.

MRESA assumes the existence of L independent signals with power spectra $S_i(f)$ and autocorrelations

$$R_{i\tau} = \int C_\tau(f) S_i(f)\, df, \tag{24}$$

where

$$C_\tau(f) = \cos 2\pi t_\tau f. \tag{25}$$

Given initial estimates $P_i(f)$ of the power spectrum of each signal S_i, and autocorrelation measurements on the sum of the signals, MRESA provides final estimates for the S_i. In particular, if the measurements R_τ^{tot} satisfy

$$R_\tau^{tot} = \sum_{i=1}^{L} \int C_\tau(f) Q_i(f)\, df, \tag{26}$$

for lags $\tau = 0, \ldots, M$, the resulting final estimates are

$$Q_i(f) = \frac{1}{\dfrac{1}{P_i(f)} + \sum_\tau \beta_\tau C_\tau(f)}, \tag{27}$$

where the β_τ are chosen so that the Q_i satisfy the autocorrelation constraints (26) [12]. Since some initial estimates may be more reliable than others, these results have been extended recently to include a frequency-dependent weight $w_i(f)$ for each initial estimate $P_i(f)$ [13]. The larger the value of $w_i(f)$, the more reliable the initial estimate $P_i(f)$ is considered to be. With the weights included, the result (27) becomes

$$Q_i(f) = \frac{1}{\dfrac{1}{P_i(f)} + \dfrac{1}{w_i(f)} \sum_\tau \beta_\tau C_\tau(f)}. \tag{28}$$

Before generalizing MRESA to include uncertain constraints, we review here some notation and results from [12] and [14]. In [12], for each of the L signals we used a discrete-spectrum approximation

$$s_i(t) = \sum_{k=1}^{N} (a_{ik} \cos 2\pi f_k t + b_{ik} \sin 2\pi f_k t)$$

$(i = 1, \ldots, L)$ with nonzero frequencies f_k, not necessarily uniformly spaced. The a_{ik} and b_{ik} were random variables with independent, zero-mean, Gaussian initial distributions. We defined random variables

$$x_{ik} = \tfrac{1}{2}(a_{ik}^2 + b_{ik}^2) \tag{29}$$

representing the power of process s_i at frequency f_k, and we described the collection of signals in terms of their joint probability density $q^\dagger(\mathbf{x})$, where $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_L)$ and $\mathbf{x}_i = (x_{i1}, \ldots, x_{iN})$. We expressed the power spectrum S as an expectation

$$S_i(f_k) = \int x_{ik} q^\dagger(\mathbf{x}) d\mathbf{x}. \tag{30}$$

In terms of initial estimates $P_{ik} = P_i(f_k)$ of $S_i(f_k)$, we wrote initial estimates p of $q^\dagger$ in the form

$$p(\mathbf{x}) = \prod_{i=1}^{L} \prod_{k=1}^{N} p_{ik}(x_{ik}) \tag{31}$$

where

$$p_{ik}(x_{ik}) = \frac{1}{P_{ik}} \exp \frac{-x_{ik}}{P_{ik}}. \tag{32}$$

The assumed Gaussian form of the initial distribution of a_{ik} and b_{ik} is equivalent to this exponential form for $p_{ik}(x_{ik})$; the coefficients were chosen to make the expectation of x_{ik} equal to P_{ik}. Using (30), we wrote a discrete-frequency form of (26) as linear constraints

$$R_r^{\text{tot}} = \sum_{i=1}^{L} \sum_{k=1}^{N} \int c_{rk} x_{ik} q^\dagger(\mathbf{x})\, d\mathbf{x} \tag{33}$$

on expectation values of $q^\dagger$, where

$$c_{rk} = C_r(f_k).$$

We obtained a final estimate q of $q^\dagger$ by minimizing the relative entropy

$$H(q, p) = \int q(\mathbf{x}) \log \frac{q(\mathbf{x})}{p(\mathbf{x})} d\mathbf{x}$$

subject to the constraints ((33) with q in place of $q^\dagger$) and the normalization condition

$$\int q(\mathbf{x})\, d\mathbf{x} = 1;$$

the result had the form

$$q(\mathbf{x}) = \prod_{i=1}^{L} \prod_{k=1}^{N} q_{ik}(x_{ik}), \tag{34}$$

where the q_{ik} were related to the final estimates

$$Q_{ik} = Q_i(f_k) = \int x_{ik} q(\mathbf{x})\, d\mathbf{x}$$

of the power spectra of the s_i by

$$q_{ik}(x_{ik}) = \frac{1}{Q_{ik}} \exp \frac{-x_{ik}}{Q_{ik}}. \tag{35}$$

This led to a discrete-frequency version of (27)

$$Q_{ik} = \frac{1}{\frac{1}{P_{ik}} + \sum_{r=0}^{M} \beta_r c_{rk}} \tag{36}$$

where the β_r had to be chosen so that

$$\sum_{i=1}^{L} \sum_{k=1}^{N} c_{rk} Q_{ik} = R_r^{\text{tot}}$$

was satisfied.

To handle uncertain constraints, we first replace (26) with a bound

$$\sum_r \left\{ \sum_i \int C_r(f) Q_i(f)\, df \; - R_r^{\text{tot}} \right\}^2 \le \varepsilon^2 \tag{37}$$

on the Euclidean norm of the error vector $\boldsymbol{v}$ given by

$$v_r = \sum_i \int C_r(f) Q_i(f)\, df \; - R_r^{\text{tot}} \tag{38}$$

We write a discrete-frequency form of (37) in terms of expec[illegible] f q:

$$\sum_{r=0}^{M} \left\{ \sum_{i=1}^{L} \sum_{k=1}^{N} \int c_{rk} x_{ik} q(\mathbf{x})\, d\mathbf{x} - R_r^{\text{tot}} \right\}^2 \le [illegible]$$

This has the form (27); by (14), minimizing relative entropy su[illegible]t to these constraints gives

$$q(\mathbf{x}) = p(\mathbf{x}) \exp\left\{ -\alpha - \sum_{r=0}^{M} \beta_r \sum_{i=1}^{L} \sum_{k=1}^{N} c_{rk} x_{ik} \right\},$$

where the β_r are to be determined so that

$$\varepsilon \frac{\beta_r}{\|\boldsymbol{\beta}\|} = \sum_{i=1}^{L} \sum_{k=1}^{N} \int c_{rk} x_{ik} q(\mathbf{x}) d\mathbf{x} - R_r^{\text{tot}} \tag{39}$$

(*cf.* (16)). Using (32), we find that q has the form (34), where $q_{ik}(x_{ik})$ is proportional to

$$\exp\left\{ \frac{-x_{ik}}{Q_{ik}} - \sum_{r=0}^{M} \beta_r \sum_{i=1}^{L} \sum_{k=1}^{N} c_{rk} x_{ik} \right\}.$$

Consequently q_{ik} is given by (35) where Q_{ik} is given by (36). Rewriting (39) in terms of Q_{ik} and passing from discrete to continuous frequencies gives

$$Q_i(f) = \frac{1}{\frac{1}{P_i(f)} + \sum_r \beta_r C_r(f)}, \tag{40}$$

where the β_r are to be determined so that

$$\varepsilon \frac{\beta_r}{\|\boldsymbol{\beta}\|} = \sum_{i=1}^{L} \int C_r(f) Q_i(f)\, df \; - R_r^{\text{tot}} \tag{41}$$

The functional form (40) of the solution with uncertain constraints is the same as the form (27) for exact constraints; the difference is in the conditions that determine the β_r: (26) for exact constraints and (41) for uncertain constraints. This is a consequence of the analogous result for probability-density estimation, noted in the previous section.

In the case of the more general constraint form

$$\sum_{rs} M_{rs} v_r v_s \le \varepsilon^2 .$$

with the error vector $\boldsymbol{v}$ as in (38), it is convenient to carry the matrix through the derivation rather than transforming the constraint functions as in (16). The result is that the final estimates again have the form (27), while the conditions (41) on the β_r are replaced by

$$\varepsilon \frac{\beta'_r}{(\boldsymbol{\beta}' \cdot \mathbf{M}\boldsymbol{\beta}')^{1/2}} = \sum_{i=1}^{L} \int G_r(f) Q_i(f)\, df - R_r^{\text{tot}} \qquad (42)$$

where

$$\boldsymbol{\beta}' = \mathbf{M}^{-1}\boldsymbol{\beta}$$

In the uncertain-constraint case, when we include weights $w_i(f)$ as in [13], the functional form of the solution becomes generalized to (28); the conditions that determine the β_r, (41) or (42), remain the same.

IV. EXAMPLE

We shall use a numerical example from [12, 13]. We define a pair of spectra, S_B and S_S, which we think of as a known "background" component and an unknown "signal" component of a total spectrum. Both are symmetric and defined in the frequency band from −0.5 to +0.5, though we plot only their positive-frequency parts. S_B is the sum of white noise with total power 5 and a peak at frequency 0.215 corresponding to a single sinusoid with total power 2. S_S consists of a peak at frequency 0.165 corresponding to a sinusoid of total power 2. Figure 1 shows a discrete-frequency approximation to the sum $S_B + S_S$, using 100 equispaced frequencies. From the sum, six autocorrelation were computed exactly. S_B itself was used as the initial estimate P_B of S_B —*i.e.*, P_B was Figure 1 without the left-hand peak. For P_S we used a uniform (flat) spectrum with the same total power as P_B. Figure 2 shows unweighted multisignal RESA final estimates Q_B and Q_S [12]. The signal peak shows up primarily in Q_S, but some evidence of it is in Q_B as well. This is reasonable since P_B, although exactly correct, is treated as an initial estimate subject to change by the data. The signal peak can be suppressed from Q_B and enhanced in Q_S by weighting the background estimate P_B heavily [13].

In Figure 3 we show final estimates for uncertain constraints with an error bound of $\varepsilon = 1$. The Euclidean distance (*i.e.*, a constraint of the form (37)) was used. The estimates were obtained with Newton-Raphson algorithms similar to those developed by Johnson [15]. Both final estimates in Figure 3 are closer to the corresponding initial estimates than is the case in Figure 2, since the sum of the final estimates is no longer constrained to satisfy the autocorrelations. Figure 4 shows results for $\varepsilon = 3$; the final estimates are even closer to the initial estimates. Because the example was constructed with exactly known autocorrelations, it is not surprising that that the exactly constrained final estimates are better than those in Figures 3 and 4 which illustrate the more conservative deviation from initial estimates that results from incorporating the uncertain constraints.

V DISCUSSION

A pleasant property of the new estimator, both in its general probability-density form and in the power-spectrum form, is that it has the same functional form as that for exact constraints. In the case of the power spectrum estimator, this means that resulting final estimates are still all-pole spectra whenever the initial estimates are all-pole and the weights are frequency-independent.

It appears that Ables was the first to suggest using an uncertain constraint of the Euclidean form (37) in MESA [16]. The use of this and a weighted Euclidean constraint in MESA was studied by Newman [17, 18]. This corresponds to a diagonal matrix **M** in (11). The generalization to general matrix constraints has been studied by Schott and McClellan [19], who offer advice on how to choose **M** appropriately. The results presented herein differ in two main respects: treatment of the multisignal case and inclusion of initial estimates. Uncertain constraints have also been used in applying maximum entropy to image processing [20, 21], although with a different entropy expression [22].

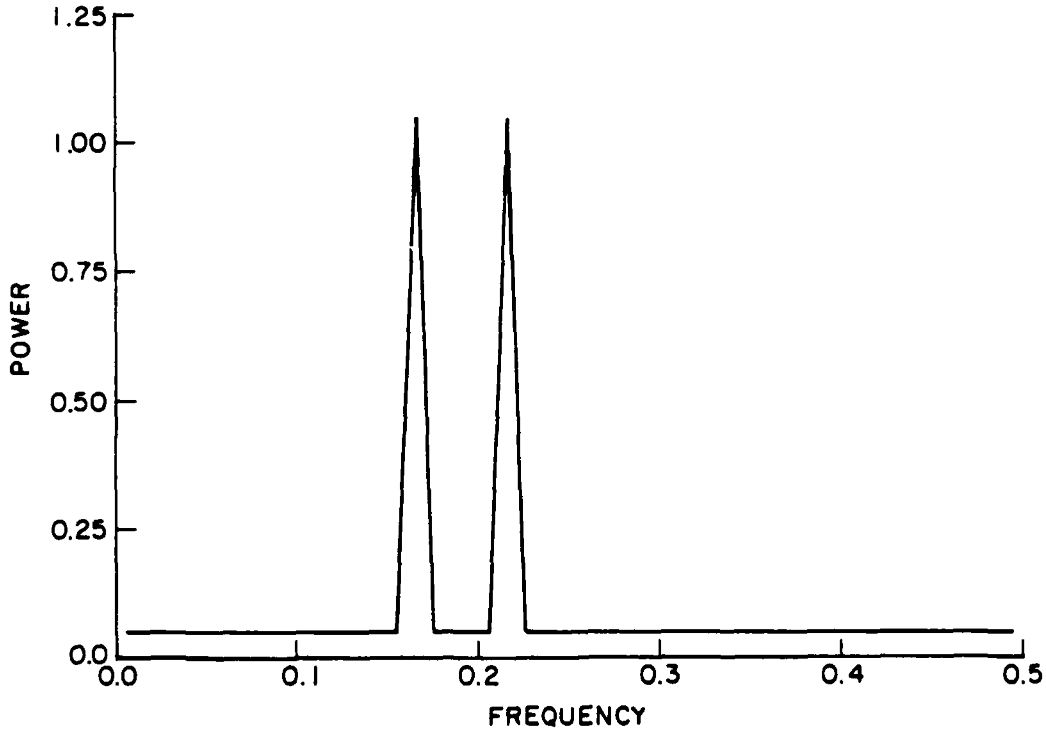


Fig. 1. Sum $S_3 + S_5$ of original spectra.

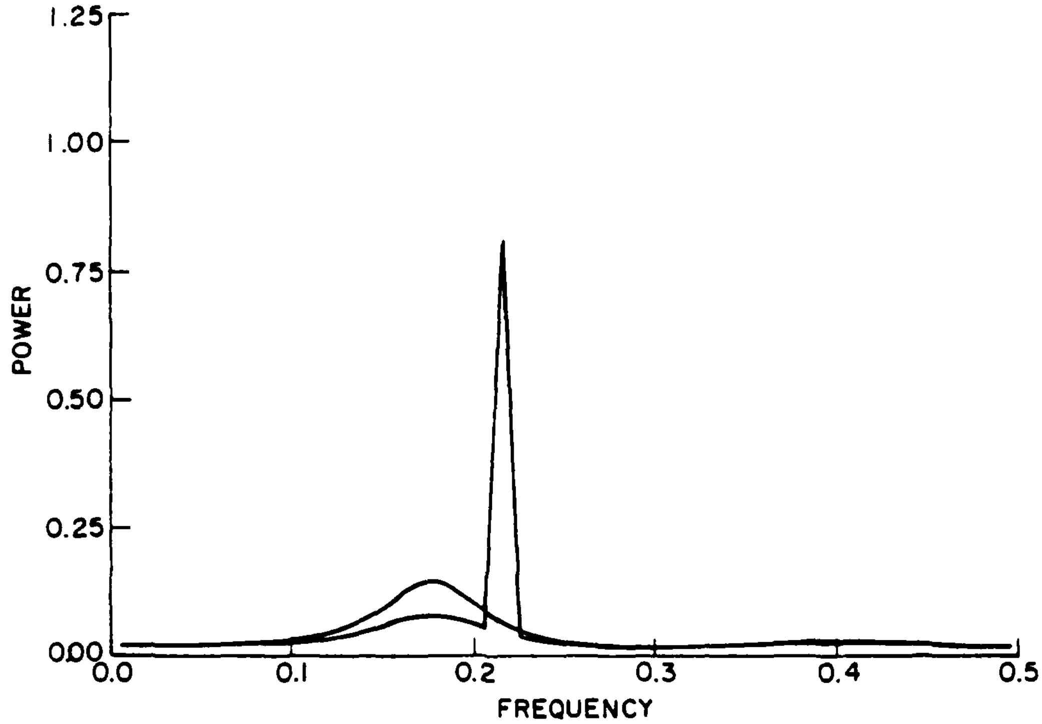


Fig. 2. MRESA final estimates Q_B and Q_S

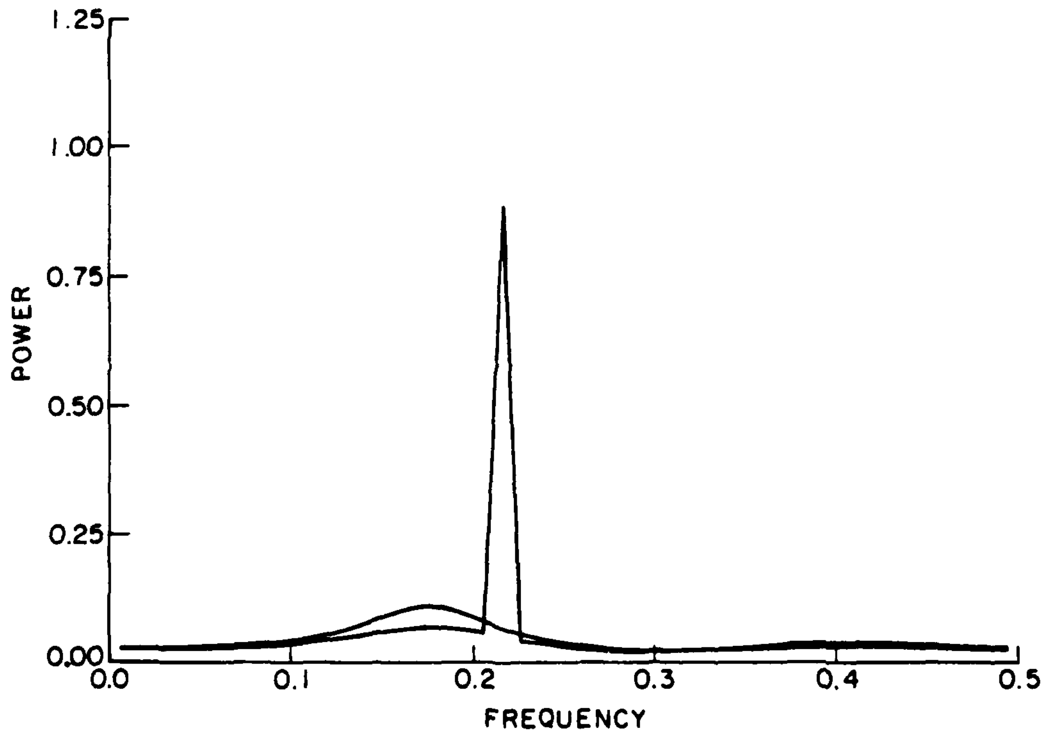


Fig. 3. Final estimates Q_B and Q_S with $\varepsilon = 1$

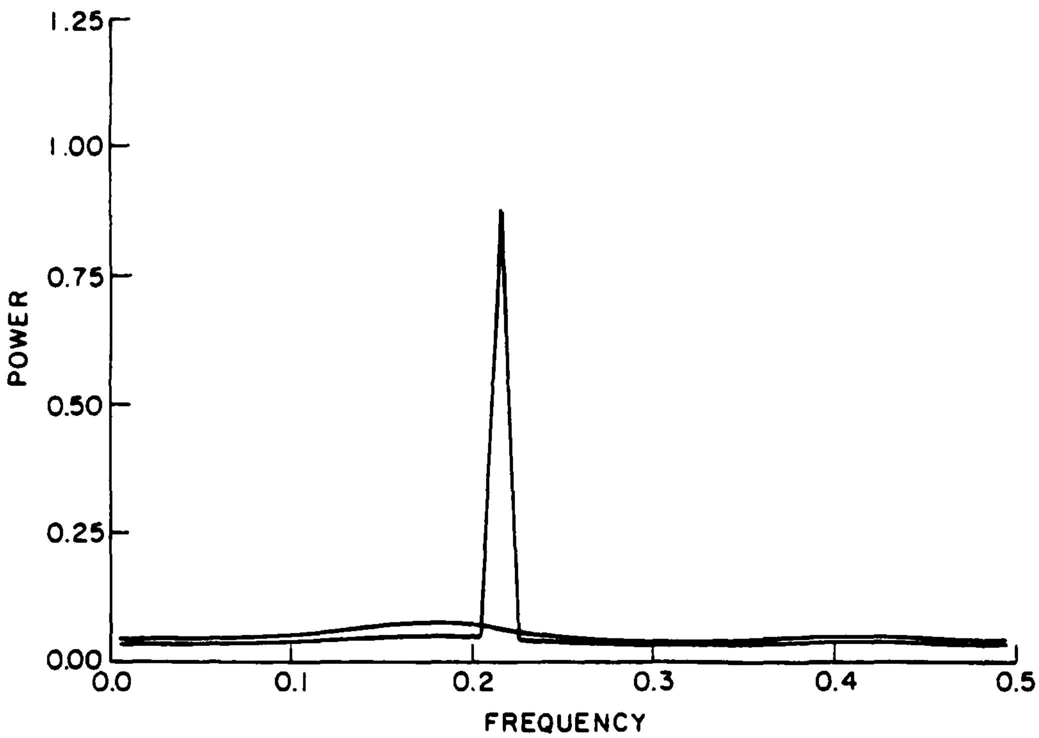


Fig. 4. Final estimates Q_B and Q_S with $\varepsilon = 3$.

REFERENCES

1. I. Csiszár, "I-divergence geometry of probability distributions and minimization problems," *Ann. Prob.* **3**, pp. 146-158 (1975).
2. J. E. Shore and R. W. Johnson, "Properties of cross-entropy minimization," *IEEE Trans. Inform. Theory* **IT-27**, pp. 472-482 (July 1981).
3. R. W. Johnson, "Determining probability distributions by maximum entropy and minimum cross-entropy," *APL79 Conference Proceedings*, pp. 24-29, ACM 0-89791-005 (May, 1979).
4. S. Kullback, *Information Theory and Statistics*, Dover, New York (1968). Wiley, New York, 1959
5. I. J. Good, "Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables," *Ann. Math. Stat.* **34**, pp. 911-934 (1963).
6. E. T. Jaynes, "Prior probabilities," *IEEE Trans. Systems Science and Cybernetics* **SSC-4**, pp. 227-241 (1968).
7. R. W. Johnson, "Axiomatic characterization of the directed divergences and their linear combinations," *IEEE Trans. Inform. Theory* **IT-25**, pp. 709-716 (Nov. 1979).

8. J. E. Shore and R. W. Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Trans. Inform. Theory* **IT-26**, pp. 26-37 (Jan. 1980). (See also comments and corrections in *IEEE Trans Inform. Theory* **IT-29**, Nov. 1983, p. 942)

9. J. P. Burg, "Maximum entropy spectral analysis," presented at the 37th Annual Meeting Soc. of Exploration Geophysicists, Oklahoma City, Okla. (1967).

10. J. P. Burg, "Maximum Entropy Spectral Analysis," Ph.D. Dissertation, Stanford University, Stanford, California (1975). (University Microfilms No. AAD75-25,499)

11. J. E. Shore, "Minimum cross-entropy spectral analysis," *IEEE Trans. Acoust., Speech, Signal Processing* **ASSP-29**, pp. 230-237 (Apr. 1981).

12. R. W. Johnson and J. E. Shore, "Minimum-cross-entropy spectral analysis of multiple signals," *IEEE Trans. Acoust., Speech, Signal Processing* **ASSP-31**, pp. 574-582 (June 1983). Also see NRL MR 4492 (AD-A097531)

13. R. W. Johnson, J. E. Shore, and J. P. Burg, "Multisignal minimum-cross-entropy spectrum analysis with weighted initial estimates," *IEEE Trans. Acoustics, Speech, Signal Processing* **ASSP-32**, pp. 531-539 (June, 1984).

14. J. E. Shore, "Minimum Cross-Entropy Spectral Analysis," NRL Memorandum Report 3921, Naval Research Laboratory, Washington, DC 20375 (Jan. 1979). (AD-A064183)

15. R. W. Johnson, "Algorithms for single-signal and multisignal minimum-cross-entropy spectral analysis," NRL Report 8667, Naval Research Laboratory, Washington, DC (August 1983). (AD-A132 400)

16. J. G. Ables, "Maximum entropy spectral analysis," *Astron. Astrophys. Suppl.* **15**, pp. 383-393 (1974).

17. W. I. Newman, "Extension to the Maximum Entropy Method," *IEEE Trans. Inform. Theory* **IT-23**, pp. 89-93 (January 1977).

18. W. I. Newman, "Extension to the Maximum Entropy Method III," *Proc. First ASSP Workshop on Spectral Estimation*, pp. 3.2.1-3.2.7 (Aug. 1981).

19. J.-P. Schott and J. H. McClellan, "Maximum entropy power spectrum estimation with uncertainty in correlation measurements," *Proc. ICASSP 83*, pp. 1068-1071, IEEE (April 1983).

20. S. F. Gull and G. J. Daniell, "Image reconstruction from incomplete and noisy data," *Nature* **272**, pp. 686-690 (April, 1978).

21. J. Skilling, "Maximum entropy and image processing — algorithms and applications," *Proc. First Maximum Entropy Workshop* (1981).

22. J. E. Shore, "Inversion as logical inference — theory and applications of maximum entropy and minimum cross-entropy," *Proceedings American Mathematical Society Symposium on Inverse Problems* (13 April 1983).

END

FILMED

2-85

DTIC